DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES

2003-05

McMASTER UNIVERSITY

Department of Economics
Kenneth Taylor Hall 426
1280 Main Street West
Hamilton, Ontario, Canada
L8S 4M4

http://socserv.socsci.mcmaster.ca/~econ/

An Efficient Mechanism for Differentiated Quality of Internet Service

Seungjin Han*
McMaster University

September 20, 2003

Abstract

The Internet industry has realized the importance of provisioning different quality of service to different applications. This paper proposes the *integrated differentiated services* that achieves the efficient throughput allocation without significant queue management or real-time pricing costs. Competing differentiated services integrates networks into a single network in an economy and allows endusers to submit packets to many networks. In equilibrium, each network posts any price for a submitted packet over time by virtue of the revenue equivalence property.

1 Introduction

As the Internet has grown into a commercial reality, the number of endusers and applications has expanded exponentially. At the beginning of the Internet, "Internet Services" meant basic electronic services such as file transfer, electronic mail, or remote login. Today, "Internet Services" refer various

*Department of Economics, Kenneth Talyor Hall, Room 426, McMaster University 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4M4. Email:hansj@mcmaster.ca. I would like to thank Martin Osborne, Hao Li, Ettore Damiano, and seminar participants at the 2003 Canadian Economin Theory Conference in Vanouver for their valuable comments. A discussion with Peter Marbach enlightened me on the architecture of internetworks. All errors are mine alone

applications including not only simple hypertext documents but also rich media such as voice over IP (VoIP) and packetized video as well as personalized content in the form of Web portals provided from the World Wide Web.[1]

As endusers use many different applications in the Internet, the importance of provisioning different quality of service to different applications has been recognized. This paper follows one of the most common approaches (elastic traffic) in the literature that uses the enduser's throughput at a time as the measure of the quality of Internet service.[2] For example, endusers who use media content must have relatively more throughput for better and quicker performance than do endusers who use simple applications such as electronic mail or hypertext document.

Current IP networks however do not offer easy ways to identify the use of packets and subsequently give them special treatment. The original TCP/IP protocol was built on the idea of equitable access to all endusers and no special treatment for anyone.[3] The internal operation of early routers used a first-in first-out (FIFO) queuing strategy. If more packets arrived than the router could handle and the queue filled up, newly arriving packets were just dropped.

A couple of considerations should be in order in provisioning different quality of service. Pricing schemes and network architectures are not independent. Feasible pricing schemes are subject to a network architecture adopted in the Internet industry. For example, the current pricing scheme is the fixed fee given an access time because the current network architecture and TCP/IP protocol cannot guarantee different throughputs to different endusers given the same access time.[4] It implies that if some network archi-

---

[1] A packet is a data unit sent over a packet-switched network. IP is the Internet Protocol that provides a connectionless best-effort delivery service of datagrams across the Internet.

[2] See Shenker (1995) for the elastic traffic. A throughput measures the number of packets transmitted through a network at a given time. When a network capacity is measured in the number of packets that the network can transmit at a given time, the term, "throughput" is used in the computer science literature instead of bandwidth.

[3] TCP stands for the Transmission Control Protocol. It is a connection-oriented transport protocol of the Internet architecture. TCP provides a reliable byte-stream delivery service.

[4] For example, unlimited access for $20 a month and 100 hours for $10 a month. If a network can assign different bandwidth for different endusers, then it is possible that given 100 hours, fast service (more bandwidth) for $15 a month and slow service for $10 a month. There have been massive requests to change the current standard network architecture

tecture and corresponding protocol are standardized, they will determine the nature of competition in the Internet industry. On the other hand, although some pricing scheme is appealing for the efficient throughput allocation, it can be quite costly to implement network architectures that support this particular pricing scheme.

MacKie-Mason and Varian (1994, 1996) suggest in a descriptive way the multiunit uniform-price auction in which endusers submit their bids for packets in real time and they pay transmitted packets at the price equal to the highest bid among dropped packets.[5] The auction approach is appealing in the sense that it generates the efficient throughput and the network supplier does not need to change the price over time. However, it is quite costly to adopt a network architecture to support the auction especially in high-speed networks because it needs a significant amount of time to calculate the market-clearing price given many endusers' bids in every instant.

The auction transmits packets at a router in the order of prices marked in the headers of submitted packets in every instant so that it generates continuously differentiated service classes.[6] In order to reduce the complicated queue management and high calculation burden to derive the market-clearing price in every instant, the scalable service classes has been proposed in the computer science literature. It categorizes submitted packets at a router into scalable service classes such as guaranteed, predictive, and best-offer. Different service classes are treated differently (see David et al. (1998) and Korilis and Lazar (1995)).

The Internet Engineering Task Force (IETF) proposed the Differentiated Services architecture on which scalable service classes can be built (see Blake et al. (1998)).[7] An important feature of the Differentiated Services archi-

---

by private network suppliers. Their requests must come from better opportunities for profit increase by changing the network architecture rather than from purely efficiency consideration.

[5]Gupta et al. (1997) consider the situation where the network dynamically changes the prices associated the different service classes in order to track a socially optimal allocation. This pricing scheme is theoretically appealing but expensive and complex to implement as Marbach (2001) points out.

[6]A router is a network node connected to two or more networks that forwards packets from one network to another.

[7]The Internet Engineering Task Force is a task force of the Internet Activity Board, responsible for providing short-term engineering solutions for the Internet. The Internet Activity Board is the main body that oversees the development and standardization of protocols of the Internet architecture. The Differentiated Services architecture is quite

tecture is that by aggregating a multitude of flows into a small number of classes, The Differentiated Services architecture eliminates the need to recognize and store information about each individual flow at core routers and it makes Differentiated Services scalable.

Computer scientists and engineers have implanted economics to see the role of pricing scalable service classes that has been proposed in terms of technical architecture. Odlyzko (1999) and Orda and Shimkin (2000) construct an economic model for scalable service classes. Fulp and Reeves (1998) and Wang and Schulerine (2001) are more interested in constructing a billing architecture to implement scalable service classes. In particular, Marbach (2001) shows that scalable service classes can achieve the efficient throughput allocation that the auction generates to network suppliers in a competitive market.

All these models focus on the endusers' choice behavior assuming that there exists a market-clearing price for each service class in every instant. It is however the network suppliers who decide their prices for different service classes. In fact, the pricing scheme for scalable service classes is quite complicated to implement. Networks adopting scalable service classes must keep changing the prices for different service classes in every instant even in a very competitive environment because the number of endusers changes in real time. Moreover, they need to keep track of the throughputs of different service classes for each enduser to decide payment.

One of the contributions of this paper is to show that there exists a pair of a simple pricing scheme and a corresponding network architecture that achieves the efficient throughput allocation without significant queue management and real-time pricing costs that scalable service classes or the multiunit uniform-price auction should bear. This paper proposes the *integrated differentiated services*, based on the Differentiated Services architecture, in the presence of finite network suppliers in an economy.

While scalable service classes or the multiunit uniform-price auction provide ways to offer different service classes within an individual network, they restrict endusers to submit their packets and pay money to a single network. The noble feature of the integrated differentiated services is that it integrates networks into a single big network in an economy. Each individual network provides a single service class possibly different from ones provided

---

a general network architecture to provide different quality of service by giving packets special handling.

by the other networks and posts a price for a submitted packet. Endusers can submit their packets to many networks. A packet is passed to the network marked on the packet header.

Each enduser faces the portfolio decision on the numbers of submitted packets between networks. In each network, packets are dropped at a router in proportion to the number of packets submitted by each enduser if the number of submitted packets is greater than the number of packets that the network can transmitted (the throughput of the network). If the number of submitted packets is smaller than the throughput of the network, then the numbers of transmitted packets for endusers are determined in proportion to the number of packets submitted by each enduser. The total number of packets submitted to a network determines the service class of the network, which is characterized by the number of transmitted packets per submitted packet. When endusers see prices for submitted packets that network suppliers post, they can foresee the market-clearing price for a transmitted packet and subsequently the unique aggregate submitted packets given the prices for submitted packets across networks. In equilibrium, each enduser chooses the portfolio such that her total throughput (the total number of her transmitted packets through all networks) coincides with the efficient throughput under the multiunit uniform-price auction, so that there is no arbitrage opportunity to endusers in equilibrium.

The most remarkable result is that the revenue equivalence property holds in any equilibrium. Regardless of prices for submitted packets that network suppliers post, endusers' portfolio decision makes each enduser's total throughput the same as the efficient one that would have been chosen under the multiunit uniform-price auction. Each enduser's payment is equal to the one that she would have paid under the multiunit uniform-price auction. The equilibrium revenue of a network is always the market-clearing price times its throughput regardless of its price for a submitted packet. The revenue equivalence property makes it possible for each network to maintain its price for a submitted packet the same over time even if the number of endusers differs at different times.

In contrast to the multiunit uniform-price auction, each network does not need to worry about whose packets are transmitted at the router in order to decide the payment since the payment is decided based on the number of submitted packets but not on the number of the transmitted packets at the router. Moreover, each network does not need to offer many service classes because the integrated differentiated services generates the same revenue that the scalable service classes would have generated. If any pricing scheme is considered for provisioning different quality of service, it should be simple enough not to delay the transmission of packets. The equilibrium price scheme described in this chapter fits in with this consideration on the top of the efficient throughput allocation.

2 Preliminaries

It is assumed that time is discrete and consists of cycles. One cycle is a set of consecutive discrete time slots, $\mathcal{T} = \{1, \cdots, T\}$. Consider a network with its fixed throughput k in each time slot $t \in \mathcal{T}$. Suppose that c is the cost of providing a network with the throughput k. For a moment, we assume that each enduser accesses the Internet at a constant and deterministic rate in each time slot of a cycle. It is called smooth traffic. We will allow bursty traffic in section 3.4.2 where the measure of total endusers changes over time slots.

There is a continuum of endusers on line whose types are parametrized by $\theta \in \Theta = [\underline{\theta}, \overline{\theta}]$. A type can be thought of as an application that an enduser uses in the Internet: for example, a simple email service, VoIP, packetized video, and etc. The measure of endusers is characterized by G, so the measure of endusers in a subset $B \subset \Theta$ is $G(B)$. When the enduser of type θ accesses the Internet in a time slot, the enduser's payoff in a time slot is

$$U(y, m, \theta) = u(y, \theta) - m \tag{1}$$

where y is the enduser's throughput (in other words, the number of her transmitted packets in a time slot) and m is the money that she pays. We assume that u is increasing in y. This utility function represents the enduser's preferences in the elastic traffic in which the enduser perceives the quality of service as an increasing function of her throughput.[8] Let $u(y, \theta)$ be *the benefit of the internet service* for the enduser of type θ when y is her throughput. u is assumed to be bounded and twice continuously differentiable. u is also assumed to be a strictly concave function with respect to y given $\theta \in \Theta$.

[8]The more packets get transmitted, the quicker an enduser can use an application. For example, an enduser can watch a real-time movie without any buffering if a stream of many packets keeps getting transmitted. This does not mean that a packet itself literally moves more quickly. The speed of the transmission cannot be higher than the speed of light.

MacKie-Mason and Varian (1996) propose the multiunit uniform-price auction for pricing the quality of service in computer networks and Crémer and Hariton (1999) examine MacKie-Mason and Varian's idea. When the network runs the multiunit uniform-price auction, each enduser submits her demand function for packets. Let $\widehat{b}(p, \theta)$ be the equilibrium demand function submitted by the enduser of type θ. The network calculates the market-clearing price $\widehat{p}(G, k)$ given demand functions submitted by endusers. It then transmits packets from the one marked with the highest price to the one marked with the market-clearing price in turn. All the packets marked with prices lower than the market-clearing prices are dropped. In this way the multiunit uniform-price auction generates continuous service classes. It is important to note that the market-clearing price is the price for a transmitted packet but not for a submitted packet. Endusers pay money only for transmitted packets but not for their submitted packets.

The payment is the market-clearing price times the throughput. While the multiunit uniform-price auction generates the efficient throughput allocation as MacKie-Mason and Varian shows, it is very costly to implement auctions for pricing the packets. At any point in time, the network must calculate the market-clearing price given submitted demand functions through sorting out the demands from the one with the highest price. It takes the network a significant time to sort out the demand functions especially in high-speed networks because the throughput is huge.

3 Differentiated Services Architecture

The Differentiated Services architecture is a network architecture for supporting different quality of service proposed by IETF. The key concept of the Differentiated Services architecture is that individual host-to-host microflows are aggregated into a single larger aggregate flow and then that single aggregate flow receives special treatment.

Currently the Differentiated Services architecture is being considered to support scalable service classes within a network. It classifies individual microflows at the edge of the network into one of several unique service classes (such as gold, silver, and bronze) and then applies a per-class service in the middle of the network. The classification is performed at the network ingress based on an analysis of one or more fields in the packet. The packet is then marked (turning on some code points, or bits, in the packet header) as belonging to a particular service class and then injected into the network.[9] The core routers that forward the packet examine the DS code points (DSCP) in the packet header to determine how the packet should be treated (for example, what transmission queue the packet should be placed in).[10]

To accomplish this, the Differentiated Services architecture defines several components. The DS-field is a bit pattern contained in the header of each packet that denotes the service (termed per-hop behavior) the packet should receive at each hop as it is forwarded through the network.[11] The per-hop behavior (PHB) defines the service the packet receives at each hop as it is forwarded through the network. A PHB may be expressed in relative (compared to other PHBs) or absolute (such as bandwidth or delay) terms. A behavior aggregate (BA) is a group of packets with the same DSCP. A PHB is applied to each BA inside the network. The boundary router is positioned at the edge of a DS-capable network. This device is responsible for packet classification, metering, marking, and possibly traffic conditioning (such as policing or shaping).[12]

4 Integrated Differentiated Services

I proposes the integrated differentiated services based on the Differentiated Service architecture, which is different from scalable service classes proposed in the computer science literature. There is a set of network suppliers, $\mathcal{I} = \{1, \cdots, I\}$, in the economy. Each network in the economy employs the Differentiated Services architecture. Even if an enduser's computer is physically connected to a single network, she can submit her packets to all the networks in the economy by marking the identity of network in the packet header. A PHB is therefore the identity of a network. While scalable service classes differentiates service classes within a single network, the integrated

[9] Marking is the process of setting the DS codepoint in a packet based on defined rules.

[10] A DS code point is a specific value of the DSCP portion of the DS field, used to select a PHB.

[11] A hop is a term used in routing. It is a path to a destination on a network is a series of hops through routers away from the origin.

[12] Metering is the process of measuring the temporal properties (e.g. rate) of a traffic stream selected by a classifier. Polishing is the process of discarding packets within a traffic stream in accordance with the state of a corresponding meter enforcing a traffic profile. Shaping is the process of delaying packets within a traffic stream to cause it to conform to some defined traffic profile.

differentiated services described in this section differentiates service classes between networks such that a single network provides only one service class possibly different from the ones provided by the other networks.

Now I show how the integrated differentiated services works technically in each network. The throughput of network i for each $i \in \mathcal{I}$ is k_i. The cost of providing a network with throughput k_i is $c_i > 0$. Suppose that $z_i(\theta)$ be the number of packets submitted to network i by the enduser of type θ in a time slot for all $\theta \in \Theta$ and all $i \in \mathcal{I}$. Then, the total number of submitted packet to network i is

$$\overline{z}_i = \int_{\underline{\theta}}^{\overline{\theta}} z_i(\theta) dG$$

The throughput of the enduser of type θ in network i is

$$z_i(\theta)\frac{k_i}{\overline{z}_i}$$

If $\frac{k_i}{\overline{z}_i} < 1$, then only $z_i(\theta)\frac{k_i}{\overline{z}_i}$ packets are transmitted for the enduser of type θ and $z_i(\theta)[1 - \frac{k_i}{\overline{z}_i}]$ packets are dropped. If $\frac{k_i}{\overline{z}_i} \geq 1$, then again $z_i(\theta)\frac{k_i}{\overline{z}_i}$ packets are transmitted for the enduser of type θ, which is more than the number of packets $z_i(\theta)$ that the enduser of type θ submits.[13]

The total throughput of all endusers in network i is

$$\int_{\underline{\theta}}^{\overline{\theta}} z_i(\theta)\frac{k_i}{\overline{z}_i} dG = k_i$$

which means that the total throughput of all endusers in network i is the same as the throughput of network i. If $[z_1, \cdots, z_I]$ is the array of the numbers of packets submitted to all networks by the enduser of type θ, then the throughput of the enduser of type θ is

$$\sum_{l=1}^{I} z_l(\theta)\frac{k_l}{\overline{z}_l}$$

Given the definition of the utility function, the benefit of the internet service for the enduser of type θ is

$$u\left(\sum_{l=1}^{I} z_l(\theta)\frac{k_l}{\overline{z}_l}, \theta\right)$$

[13]Note that a stream of packets travels in networks. Even if very few packets are submitted at a router in the case of $\frac{k_i}{\overline{z}_i} \geq 1$, more packets can be transmitted instantly.

4.1 Smooth Traffic

Network suppliers' pricing strategies and endusers' purchase decisions are as follows. First, network supplier i for each $i \in \mathcal{I}$ posts a price for a submitted packet $p_i \in \mathcal{R}_+$. Second, each enduser decides how many packets to submit to which networks after seeing prices for submitted packets $\mathbf{p} = (p_1, \cdots, p_I) \in \mathcal{R}_+^I$ that network suppliers post.

A purchase decision is characterized by a mapping $s : \mathcal{R}_+^I \times \Theta \to \mathcal{R}^I$. $s(\mathbf{p}, \theta) = (s_1(\mathbf{p}, \theta), \cdots, s_I(\mathbf{p}, \theta))$ is an array of the numbers of packets submitted by the enduser of type θ when $\mathbf{p} = (p_1, \cdots, p_I)$ is prices for submitted packets that network suppliers post. Let $\overline{x}_i(\mathbf{p}, s)$ be the total number of packets submitted to network i when prices for submitted packets are $\mathbf{p}$ and the purchase decision chosen by endusers is s. It is expressed by

$$\overline{x}_i(\mathbf{p}, s) = \int_{\underline{\theta}}^{\overline{\theta}} s_i(\mathbf{p}, \theta) dG \tag{2}$$

The throughput of the enduser of type θ is

$$b(\mathbf{p}, \theta) = \sum_{i=1}^{I} s_i(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s)} \tag{3}$$

Suppose that s is the purchase decision chosen by endusers. When prices for submitted packets are $\mathbf{p}$, the payoff for the enduser of type θ is

$$V(s, \mathbf{p}, \theta) = u\left(\sum_{i=1}^{I} s_i(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s)}, \theta\right) - \sum_{i=1}^{I} p_i s_i(\mathbf{p}, \theta) \tag{4}$$

A purchase decision s^* is a continuation equilibrium if s^* satisfies

$$s^*(\mathbf{p}, \theta) \in \underset{(x_1, \cdots, x_I) \in \mathcal{R}^I}{\arg\max}\ u\left(\sum_{i=1}^{I} x_i \frac{k_i}{\overline{x}_i(\mathbf{p}, s)}, \theta\right) - \sum_{i=1}^{I} p_i x_i \tag{5}$$

for every $\mathbf{p} \in \mathcal{R}_+^I$ and every $\theta \in \Theta$, where $(x_1, \cdots, x_I)$ is the numbers of submitted packets to all networks (x_i is the number of packets submitted to network i for all $i \in \mathcal{I}$)

Suppose that network supplier i uses a pricing strategy $p_i \in \mathcal{R}_+$ (a price for a packet submitted to network i). Let $\mathbf{p}_{-i} = (p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_I)$. Given a continuation equilibrium s^*, the payoff for network supplier $i \in \mathcal{I}$ is

$$\Pi(p_i, \mathbf{p}_{-i}, s^*) = p_i \overline{x}_i(\mathbf{p}, s^*) - c_i \tag{6}$$

where $\mathbf{p} = (p_1, \cdots, p_I) \in \mathcal{R}_+^I$.

An *integrated differentiated services equilibrium* is an array of pricing strategies $\mathbf{p}^* = (p_1^*, \cdots, p_I^*)$ and a continuation equilibrium s^* such that $\mathbf{p}^*$ is a Nash equilibrium for the normal form game defined by the continuation equilibrium s^*.

First, we derive the necessary condition that should be satisfied in any continuation equilibrium. This necessary condition is surprising because any continuation equilibrium reproduces the equilibrium allocation that would have been achieved under the multiunit uniform-price auction.

Theorem 1 *In any continuation equilibrium s^*, the following conditions are satisfied.*

1. for every $\theta \in \Theta$ and every $\mathbf{p} = (p_1, \cdots, p_I) \in \mathcal{R}_+^I$,

$$V(s^*, \mathbf{p}, \theta) = \tag{7}$$

$$u\left(\widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right), \theta\right) - \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right)$$

2. Revenue Equivalence Property: for every $i \in \mathcal{I}$ and every $\mathbf{p} = (p_1, \cdots, p_I) \in \mathcal{R}_+^I$

$$p_i \overline{x}_i(\mathbf{p}, s^*) = \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i \tag{8}$$

Proof. See Appendix 6.1.

Suppose that s^* characterizes the endusers' equilibrium purchase decision. For every $\theta \in \Theta$, every $\mathbf{p} \in \mathcal{R}_+^I$, $s^*(\mathbf{p}, \theta) = (s_1^*(\mathbf{p}, \theta), \cdots, s_I^*(\mathbf{p}, \theta))$ must satisfy (5). The first-order condition is

$$\frac{\partial u}{\partial y}\left(\sum_{i=1}^{I} s_i^*(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}, \theta\right) = \frac{p_i}{\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}} \tag{9}$$

for every $i \in \mathcal{I}$. The throughput of the enduser of type θ in network i is $s_i^*(\mathbf{p}, \theta) k_i / \overline{x}_i(\mathbf{p}, s^*)$. In equilibrium, the aggregate demanded throughput in each network is equal to the throughput of the network: $\int [s_i^*(\mathbf{p}, \theta) k_i / \overline{x}_i(\mathbf{p}, s^*)] dG =$

k_i for all $i \in \mathcal{I}$. The price for a transmitted packet is $p_i \overline{x}_i(\mathbf{p}, s^*)/k_i$. Since it is the price for a transmitted packet that clears the market in the economy, it must be equal to the market-clearing price under the multiunit uniform-price auction.

$$\frac{p_i \overline{x}_i(\mathbf{p}, s^*)}{k_i} = \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \tag{10}$$

where $\widehat{p}(G, \sum_{l=1}^{I} k_l)$ is the market-clearing price when the measure of endusers is G and the total throughput in the economy is $\sum_{l=1}^{I} k_l$. Under this price for a transmitted packet, the total throughput of the enduser of type θ is equal to the throughput she would have purchased under the multiunit uniform-price auction. The payment made by the enduser of type θ is

$$\begin{aligned} \sum_{i=1}^{I} p_i s_i^*(\mathbf{p}, \theta) &= \sum_{i=1}^{I} \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) s_i^*(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} \\ &= \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) \end{aligned} \tag{11}$$

where $\widehat{b}(\widehat{p}(G, \sum_{l=1}^{I} k_l), \theta)$ is the throughput she would have purchased under the multiunit uniform-price auction. That is, the equilibrium payment is the same as the payment that she would have made for $\widehat{b}(\widehat{p}(G, \sum_{l=1}^{I} k_l), \theta)$ under the multiunit uniform-price auction. Therefore, every enduser will have the same payoff as she would obtain under the multiunit uniform-price auction. Moreover, it is obvious from (10) that each network has the revenue that is equal to its throughput times the market-clearing price no matter what price for a submitted packet a network supplier posts: $p_i \overline{x}_i(\mathbf{p}, s^*) = \widehat{p}(G, \sum_{l=1}^{I} k_l) k_i$ for all $i \in \mathcal{I}$.

Given the price for a submitted packet that network i posts, the equilibrium aggregate submitted packets make the price for a transmitted packet the same as the market-clearing price: the higher price for a submitted packet a network supplier posts, the less packets are submitted. Therefore, there is no arbitrage opportunity by choosing any particular network in equilibrium. The equilibrium throughput allocation is efficient since it is exactly the same as the one under the multiunit uniform-price auction. The existence of equilibria is not yet addressed. Theorem 2 identifies an integrated differentiated services equilibrium.

Theorem 2 $(\mathbf{p}^*, s^*)$ *is an integrated differentiated services equilibrium if* p_i^* *is any positive price for a submitted packet for every* $i \in \mathcal{I}$ *and* $s^*(\mathbf{p}, \theta) = (s_1^*(\mathbf{p}, \theta), \cdots, s_I^*(\mathbf{p}, \theta))$ *satisfies, for every* $\mathbf{p} \in \mathcal{R}_+^I$*, every* $\theta \in \Theta$*,*

$$s_i^*(\mathbf{p}, \theta) = \frac{\widehat{p}\left(G, \sum_{l=1}^I k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^I k_l\right), \theta\right) k_i}{p_i \sum_{l=1}^I k_l} \tag{12}$$

Proof. See Appendix 6.2.

No matter what price for a submitted packet network i posts, its revenue is always equal to $\widehat{p}(G, \sum_{l=1}^I k_l) k_i$ in equilibrium. It is therefore obvious that each network supplier's equilibrium pricing strategy is to post any positive price for a submitted packet. To show the existence of an equilibrium, we must find the endusers' equilibrium purchase decision and then show the sum of endusers' packets submitted to network $i \in \mathcal{I}$ is equal to $\widehat{p}(G, \sum_{l=1}^I k_l) k_i / p_i$. Suppose that the purchase decision is the one in theorem 2. With this purchase decision, each enduser submits packets to network i equal to $\widehat{b}(\widehat{p}(G, \sum_{l=1}^I k_l), \theta) \overline{x}_i(\mathbf{p}, s^*) / \sum_{l=1}^I k_l$. Then the enduser's total throughput is equal to the throughput that she would have purchased under the multiunit uniform-price auction as appendix 6.2 proves. It is straightforward to show that in fact integrating endusers' packets submitted to network i yields the aggregate number of submitted packets such that it makes the price for a transmitted packet equal to the market-clearing price. Therefore, $(\mathbf{p}^*, s^*)$ described in theorem 2 is an integrated differentiated services equilibrium.

4.2 Bursty Traffic

The integrated differentiated services described in the previous section focuses on the smooth traffic where each enduser accesses the Internet at a constant and deterministic rate. In practice, the measure of endusers who access the Internet does vary over time slots. It is called bursty traffic. Traffic in the Internet can be very congested in some time slot relative to other time slots. First, it is interesting to know whether or not the integrated differentiated services can induce the efficient throughput allocation in bursty traffic. Second, we like to know whether or not there exists a pricing scheme simple enough to avoid complicated queue management or high real-time pricing costs in bursty traffic. Following the approach of Marbach (2001), we assume that in time slot $t \in \mathcal{T}$, each enduser accesses the Internet with probability q_t and stays off line with probability $1 - q_t$. The value of q_t for all $t \in \mathcal{T}$ is public information. The measure of endusers on line whose type belongs to a subset $A \subset \Theta$ is $q_t G(A)$. Therefore, the measure of endusers at time slot t is $q_t G$.

In each time slot, network suppliers post their prices for submitted packets. These prices can vary over time slots depending on network suppliers' decisions. After seeing the prices for submitted packets that network suppliers post in each time slot, endusers make purchase decisions on packets submitted across networks when they access the Internet. It is quite straightforward to derive endusers' equilibrium purchase decisions in each time slot. In smooth traffic, endusers' equilibrium purchase decision is the same over time slots because the measure of endusers on line is constant. While the measure of endusers on line varies over time slots, endusers can forecast the changes systematically because they know how heavy the traffic, $q_t G$, is in each time slot.

In each time slot $t \in \mathcal{T}$, every enduser can base her equilibrium purchase decision on the decision rules described in theorem 2 by replacing the measure of endusers on line with $q_t G$ in each time slot $t \in \mathcal{T}$. Each enduser's equilibrium purchase decision follows, for every $t \in \mathcal{T}$, every $\theta \in \Theta$, and every $\mathbf{p}_t \in \mathcal{R}_+^I$

$$s_{it}^*(\mathbf{p}_t, \theta) = \frac{\widehat{p}\left(q_t G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(q_t G, \sum_{l=1}^{I} k_l\right), \theta\right) k_i}{p_{it} \sum_{l=1}^{I} k_l} \tag{13}$$

where $\mathbf{p}_t = (p_{1t}, \cdots, p_{It})$ is the prices for submitted packets that network suppliers post in time slot t and $s_{it}^*(\mathbf{p}_t, \theta)$ is the number of packets submitted to network i by the enduser of type θ in time slot t.

The network suppliers should decide pricing strategies on how to decide the price for a submitted packet over time slots. The revenue equivalence property makes it possible for network suppliers to post the same price for a submitted packet over time slots in equilibrium as theorem 3 proves.

Theorem 3 *In equilibrium, network suppliers do not need to change their prices for submitted packets over time slots.*

Proof. See Appendix 6.3.

Given prices for submitted packets, it is the endusers' purchase decision that makes the price for a transmitted packet the same as the market-clearing price in every time slot. Since any price for a submitted packet is an equilibrium price for a submitted packet in any time slot by theorem 2, the network suppliers do not need to change their prices for submitted packets over time slots.

5 Conclusion

This paper proposes an alternative selling mechanism for the auction in a highly dynamic environment in which the object is multiunit and its dilivery time characteriszes its quality. When endusers perceive quality of the object as how quckly the object can be delivered, the auction leaves a significant pricing burden with the networks because it takes a significant amount of time to calculate the equilibrium price given instant bids.

The integrated differentiated services alleviates the network's burden of queue management by letting endusers decide the price for a transmitted packet. Networks do not need to change prices for submitted packets because of the revenue equivalence property. The environment in the Internet is quite complicated in the sense that there are many endusers accessing the Internet at the same time and the number of endusers who access the Internet changes over time even during a day. The transmission time in the Internet is just few seconds or usually less than a second. Especially, the network with high speed has a lot shorter transmission time. If any pricing scheme is considered for provisioning different quality of service, it should be simple enough not to delay the transmission of packets. The equilibrium prices described in this paper fits in this consideration on the top of the efficient throughput allocation in the sense that each network supplier posts any positive price and does not need to change it over time.

The integrated differentiated services allows the enduser to submit packets and pay money to many different networks. Even though networks do not have different service classes, posting different prices for submitted packets yields the different throughput per submitted packet across networks in the economy. Since the Differentiated Service architecture can integrate the networks in the economy into a single network and the individual networks provide different service classes by posting different prices for submitted prices, endusers can choose many different networks to submit their packets.

A network architecture restricts the set of available pricing schemes. Therefore, it is very important to come up with a network architecture that can provide an easy implementation of a pricing scheme and the efficient throughput allocation over time. The integrates differentiated services proposed in this chapter satisfies this double object.

6 Appendix

6.1 Proof of Theorem 1

Suppose that s^* is the equilibrium purchase decision for endusers. For every $\theta \in \Theta$, every $\mathbf{p} \in \mathcal{R}_+^I$, $s^*(\mathbf{p}, \theta) = (s_1^*(\mathbf{p}, \theta), \cdots, s_I^*(\mathbf{p}, \theta))$ satisfies (5). The first-order condition is, for all $i \in \mathcal{I}$, all $\mathbf{p} \in \mathcal{R}_+^I$, and all $\theta \in \Theta$,

$$\frac{\partial u}{\partial y}\left(\sum_{i=1}^{I} s_i^*(\mathbf{p}, \theta)\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}, \theta\right)\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} = p_i \tag{14}$$

Therefore,

$$\frac{\partial u}{\partial y}\left(\sum_{i=1}^{I} s_i^*(\mathbf{p}, \theta)\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}, \theta\right) = \frac{p_i}{\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}} \tag{15}$$

The second-order condition is satisfied because of the concavity assumption on u with respect to y given $\theta \in \Theta$. The throughput of the enduser of type θ in the network i is $s_i^*(\mathbf{p}, \theta)k_i/\overline{x}_i(\mathbf{p}, s^*)$ when the prices for submitted packets are $\mathbf{p}$. In equilibrium, the aggregate demanded throughput in each network i is equal to the throughput of the network:

$$\int_{\underline{\theta}}^{\overline{\theta}} s_i^*(\mathbf{p}, \theta)\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} dG = k_i \tag{16}$$

Since the right-hand side of (15) is the price for a transmitted packet, it must be equal to the market-clearing price under the multiunit uniform-price auction.

$$\frac{p_i}{\frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}} = \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \tag{17}$$

where $\widehat{p}(G, \sum_{l=1}^{I} k_l)$ is the market-clearing price when the measure of endusers is G and the total throughput in the economy is $\sum_{l=1}^{I} k_l$. Under the price for a transmitted packet, the throughput of the enduser of type θ is equal to the throughput that she would have purchased under the multiunit uniform-price auction.

$$\widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) = \sum_{i=1}^{I} s_i^*(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} \tag{18}$$

The equilibrium payoff of the enduser of type θ for every $\theta \in \Theta$ is

$$\begin{aligned} V(s^*, \mathbf{p}, \theta) &= u\left(\sum_{i=1}^{I} s_i^*(\mathbf{p}, \theta) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)}, \theta\right) - \sum_{i=1}^{I} p_i s_i^*(\mathbf{p}, \theta) \\ &= u\left(\widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right), \theta\right) - \sum_{i=1}^{I} p_i s_i^*(\mathbf{p}, \theta) \\ & u\left(\widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right), \theta\right) - \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) \end{aligned} \tag{19}$$

The first equality comes from the definition of the equilibrium purchase decision and the second equality comes from (18). From (17),

$$p_i = \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} \tag{20}$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{I} p_i s_i^*(\mathbf{p}, \theta) &= \sum_{i=1}^{I} \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} s_i^*(\mathbf{p}, \theta) \\ &= \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) \end{aligned} \tag{21}$$

The equilibrium payment is the same as the payment that she would have made for $\widehat{b}(\widehat{p}(G, \sum_{l=1}^{I} k_l), \theta)$ under the multiunit uniform-price auction. The third equality in (19) comes from (21).

From (17),

$$p_i \overline{x}_i(\mathbf{p}, s^*) = \widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i \tag{22}$$

for every $i \in \mathcal{I}$. It says that the revenue of the network i with any price for a submitted packet $p_i \in \mathcal{R}_+$ is always equal to the revenue that he would have got under the multiunit uniform-price auction.

6.2 Proof of Theorem 2

No matter what price for a submitted packet a network supplier posts, his revenue is always equal to $\widehat{p}(G, \sum_{l=1}^{I} k_l)k_i$ in equilibrium. It is therefore obvious that each network supplier's equilibrium pricing strategy is to post any positive price for a submitted packet. From (17), we know the aggregate packets submitted to network i is equal to

$$\overline{x}_i(\mathbf{p}, s^*) = \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i}{p_i} \tag{23}$$

To show the existence of an equilibrium, we must find the endusers' equilibrium purchase decision and then show the sum of individual packets submitted by each enduser in network i is equal to $\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i/p_i$.

Suppose that the equilibrium purchase strategy is, for every $\theta \in \Theta$, every $\mathbf{p} \in \mathcal{R}_+^I$,

$$s_i^*(\mathbf{p}, \theta) = \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) k_i}{p_i \sum_{l=1}^{I} k_l} \tag{24}$$

Then the enduser's total throughput is

$$\begin{aligned}
&\sum_{i=1}^{I} \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) k_i}{p_i \sum_{l=1}^{I} k_l} \frac{k_i}{\overline{x}_i(\mathbf{p}, s^*)} \qquad (25)\\
&= \sum_{i=1}^{I} \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) k_i}{p_i \sum_{l=1}^{I} k_l} \frac{p_i}{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right)}\\
&= \frac{\widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) \sum_{i=1}^{I} k_i}{\sum_{l=1}^{I} k_l}\\
&= \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right)
\end{aligned}$$

So, it is equal to the throughput that she would have purchased under the multiunit uniform-price auction. $s^*(\mathbf{p}, \theta) = (s_1^*(\mathbf{p}, \theta), \cdots, s_I^*(\mathbf{p}, \theta))$ satisfies the first-order condition in (15). Therefore, it is the enduser's equilibrium purchase decision. To get the aggregate packets submitted to network i, we integrate (24) with respect to θ.

$$\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} s_i^*(\mathbf{p},\theta)dG \qquad (26)\\
&= \int_{\underline{\theta}}^{\overline{\theta}} \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) k_i}{p_i \sum_{l=1}^{I} k_l} dG \\
&= \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i}{p_i \sum_{l=1}^{I} k_l} \int_{\underline{\theta}}^{\overline{\theta}} \widehat{b}\left(\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right), \theta\right) dG \\
&= \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i}{p_i \sum_{l=1}^{I} k_l} \sum_{l=1}^{I} k_l \\
&= \frac{\widehat{p}\left(G, \sum_{l=1}^{I} k_l\right) k_i}{p_i}
\end{aligned}$$

This is equal to the aggregate submitted packets derived from (17). Therefore, $(\mathbf{p}^*, s^*)$ described in theorem 2 is an integrated differentiated services equilibrium.

6.3 Proof of Theorem 3

Suppose that the prices for submitted packets that network suppliers are $\mathbf{p}_t$ in time slot t. In time slot t, the equilibrium price for a transmitted packet is, for all $i \in \mathcal{I}$,

$$\frac{p_{it}}{\frac{k_i}{\overline{x}_{it}(\mathbf{p}, s_t^*)}} = \widehat{p}\left(q_t G, \sum_{l=1}^{I} k_l\right) \qquad (27)$$

where $\overline{x}_{it}(\mathbf{p}, s_t^*)$ is the aggregate number of packets submitted to the network i in time slot t. Given the prices for submitted prices, the individual purchase decision makes aggregate packets submitted to network i equalize the price for a transmitted packet to the market-clearing price under the multiunit uniform-price auction. Since any price for a submitted packet is an equilibrium price for a submitted price for each network i in any given time slot by theorem 2, the network suppliers do not need to change their prices for submitted packets over time slots.

References

[1] Blake, S.; Black D., Carlson M., Davies E., Wang Z., and Weiss, W., "An Ar-chitecture for Differentiated Services," Internet Engineering Task Force, http://www.ietf.org/rfc/rfc2475.txt, 1998.

[2] Chen, Shaogong; Park, Kihong and Sitharam, Meera, "On the ordering prop-erties of GPS routers for multi-class QoS provision," Proceedings of the SPIE Inter-national Conference on Performance and Control of Network Systems, 1998, pp. 252-265.

[3] Clark, David D.; Shenker, Scott and Zhang, Lixia, "Supporting Real-time Appli-cations in an Integrated Service Packet Network: Architecture and Mechanism," Conference Proceedings Communications Archetectures & Protocols-SIGCOMM 1992, 1992, pp. 14-26

[4] Crémer, Jacques and Hariton, Cyril, "The Pricing of Critical Applications in the Internet," Journal of the Japanese and International Economics, December 1999, 13 (4), pp. 281-310.

[5] Douglis, Fred and Kaasheok, Frans M., "Scalable Internet Services," IEEE Internet Computing, July/August 2001, 5 (4), pp. 36-37.

[6] Fulp, Errin W. and Reeves, Douglas S., "Distribution Network Flow Control Based on Dynamic Competitive Markets," Proceedings of the IEEE International Conference on Network Protocol '1998, 1998, pp. 119-128.

[7] Gupta, Alok; Stahl, Dale O. and Whinston, Andrew B., "A Stochastic Equilib-rium Model of Internet Pricing," Journal of Economics Dynamics and Control, May 1997, 21 (4-5), pp. 697-722.

[8] Korilis, Yannis A. and Lazar, Aurel A., "On the Existence of Equilibria in Nonco-operative Optimal Flow Control," Journal of the ACM, May 1995, 42(3), pp. 584-613

[9] MacKie-Mason, Jeffrey K. and Varian Hal R., "Pricing the internet," in Kahin, Brian and Keller, James eds., Public Access to the Internet. Cambridge: MIT Press, 1994, pp. 269-314.

[10] MacKie-Mason, Jeffrey K. and Varian Hal R., "Some Economics of the Inter-net," in Sichel, Werner and Alexander, Donald L. eds., Netweorks, Infrastructure, and the New Task for Regulation. Ann Arbor: University of Michigan Press, 1996, pp. 107-136.

[11] Marbach, Peter, "Pricing Differentiated Services Networks: Bursty Traffic," IEEE INFOCOM 2001, 2001, (Proceedings), 2, pp. 650-658.

[12] Odlyzko, Andrew, "Paris Metro Pricing for the Internet," Proceedings of the ACM Conference on Electronic Commerce '99, 1999, pp. 140-147.

[13] Orda, Ariel and Shimkin, Nahum, "Incentive Pricing in Multiclass Systems," Telecommunication Systems, 2000, 13 (2), pp. 241-267.

[14] Wang, Xin and Schulzrinne, Henning, "Pricing Network Resources for Adaptive Applications in a Differentiated Services Network," IEEE INFOCOM 2001, 2001, (Proceedings), 2, pp.943-952.